AFRL-HE-WP-TR-2002-0050

UNITED STATES AIR FORCE RESEARCH LABORATORY

Near-Field Localization Response Methods

Douglas S. Brungart

AIR FORCE RESEARCH LABORATORY

William M. Rabinowitz

BOSE CORPORATION
Boston MA

Nathaniel I. Durlach

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

April 1998

Final Report for the Period September 1995 to March 1998

20020905 031

Human Effectiveness Directorate
Crew System Interface Division
2610 Seventh Street
Wright-Patterson AFB OH 45433-7901

Approved for public release; distribution is unlimited.

NOTICES

TECHNICAL REVIEW AND APPROVAL

AFRL-HE-WP-TR-2002-0050

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public.

The voluntary informed consent of the subjects used in this research was obtained as required by Air Force Instruction 40-402.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

MARIS M. VIKMANIS
Chief, Crew System Interface Division
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 1998	Final - September 1995 to March 1998

4. TITLE AND SUBTITLE

Near-Field Localization Response Methods

5. FUNDING NUMBERS

PE - 62202F
PR - 7184
TA - 718441
WU - 71844104

6. AUTHOR(S)

Douglas S. Brungart (Air Force Research Laboratory)
William M. Rabinowitz (Bose Corp.)
Nathaniel I. Durlach (Massachusetts Institute of Technology)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory, Human Effectiveness Directorate
Crew System Interface Division
Aural Displays and Bioacoustics Branch
Air Force Materiel Command
Wright-Patterson AFB OH 45433-7901

8. PERFORMING ORGANIZATION REPORT NUMBER

AFRL-HE-WP-TR-2002-0050

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Four response methods that allow subjects to indicate perceived locations within one meter of their heads were evaluated experimentally. In the Direct-Location (DL) method, the subject moved a response pointer directly to the perceived target location. In the Large-Head (LH) method, the subject moved the response pointer to the perceived location, relative to a manikin head, that corresponded to the location of the target relative to their own head. The Small-Head (SH) method was similar to LH, except that a half-scale manikin head was used and the subjects were asked to scale down their response distances by a factor of two. In the Verbal Report (VR) response, subjects verbally indicated the spherical coordinates of the target location. Measurements with a visual target indicated that DL was relatively unbiased and considerably more accurate than the other three methods. The three indirect methods, relatively LH, SH, and VR, were all roughly equivalent in performance. Correcting for bias improved accuracy in the LH, SH, and VR response, but even with bias correction these methods were inferior to the uncorrected DL response. When the visual target was replaced with an acoustic stimulus, the errors in the DL response were approximately doubled. The errors in the acoustic experiment were, however, roughly equivalent in the front and rear hemispheres, despite the expected difficulties of reaching behind the body and outside the visual field. The results suggest that DL is the most appropriate response method for near-field auditory localization experiments.

14. SUBJECT TERMS	15. NUMBER OF PAGES
virtual audio displays, head-related transfer funtions, auditory localization	48
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

THIS PAGE INTENTIONALLY LEFT BLANK

PREFACE

The work described herein was performed at the Noise and Vibration Branch, Force Survivability and Logistics Division, Human Effectiveness Directorate of the Air Force Research Laboratory and the Research Laboratory of Electronics at the Massachusetts Institute of Technology. At the time the research was performed, Douglas S. Brungart was a participant in the Palace Knight civilian training program at MIT, and the work described was incorporated in his doctoral dissertation (Brungart, 1998). Nathaniel Durlach and William Rabinowitz were Research Scientists at the Sensory Communication group of the Research Laboratory of Electronics and were serving on Dr. Brungart's graduate committee. Additional support was provided by AFOSR grant F49620-96-1-0202.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

In traditional psychoacoustic localization studies, subjects have been asked to estimate either the direction of a sound source or its distance. The perceived distance and direction of a sound source have been simultaneously measured only in a few limited cases. In planning an experiment to examine localization ability for nearby sound sources (less than 1 m away from the listener), a response method was required that would accurately measure localization accuracy in azimuth, elevation, and distance. Since some targets would be extremely close to the subject (5 cm or less from the head), a subject-based coordinate system was used to provide a consistent basis for expressing locations relative to the head in terms of spherical coordinates. The subject would be required to respond to source locations behind the head and out of the visual field, so a response method that was consistent for locations in front and behind the listener was desired. Furthermore, as a large number of trials was required, he response method had to be relatively fast. No response method found in the literature seemed adequate.

Four response methods were tested: Direct-Location (DL), Large-Head Transformation (LH), Small-Head Transformation (SH), and Verbal Report (VR). In the DL method, the subject moves an electromagnetic position sensor directly to the perceived location. In the LH and SH paradigms, the subject moves the position sensor to a location relative to a full-size or half-size manikin head that matches the perceived source location relative to his or her own head. In the VR response, the subject simply states the coordinates of the perceived sound in degrees azimuth, degrees elevation, and distance. These four response methods were tested by determining the accuracy of subject responses when the stimulus was a visual target. The results show that DL is superior to the other three methods, and that LH, SH, and VR are roughly comparable.

2.0 BACKGROUND

In directional localization experiments, researchers have depended on two types of response methods— verbal report and pointing. Wightman and Kistler (1992) used verbal reports to collect azimuth and elevation responses. Subjects stated the perceived source azimuth and elevation in degrees and the investigator typed these coordinates directly into a computer. This procedure has two major drawbacks. First, the verbal response must be correctly interpreted and properly entered by the investigator; as a result, this procedure has a higher chance of error than an automated system where the control computer can directly read the subject's response. Second, the response method is slow; Wightman and Kistler collected only 2-3 responses per minute using this technique.

Makous and Middlebrooks (1990) used a head-pointing response method in their directional localization experiments. Subjects wore an electromagnetic head-tracking sensor, and responded by pointing their nose in the perceived direction of the source. This method is slightly faster than the verbal report method (3-4 responses per minute) and eliminates data entry errors. However, head-pointing is difficult for locations behind and above the subject, and it complicates methods for immobilizing of the subject's head.

Gilkey and colleagues (1995) examined an alternative to these methods called the God's Eye Location Pointing (GELP) method. The subject was seated with a plastic sphere (20 cm in diameter) located 22 cm in front of the subject and approximately 50 cm below ear level. The subject moved an electromagnetic sensor on the surface of the plastic sphere to the perceived direction of the sound. This method eliminated the need for the subject to speak or move his or her head, and permitted the use of a bite bar to restrict head motion. Furthermore, the subjects never had to move their hands away from the response sphere, so they responses could be made quickly (16-20 per minute). Gilkey assessed the GELP method with two experiments. First, subjects were asked to identify the directions of sound sources; the average angle error was 18.2°. This performance is comparable to the 20° average error reported by Wightman and Kistler using verbal reports. The mean errors in azimuth and elevation were slightly lower than those reported by Makous and Middlebrooks using head-pointing. In a second assessment, the sound source was eliminated; the experimenter simply read verbal coordinates to the subject, who was then asked to respond at that location with

the GELP method. This produced mean angular errors of approximately 9° (vs. 20° with the sound source). One concern about the GELP system is the possibility that the 20-cm rigid sphere might generate unwanted reflections that confound the auditory experiment. Overall, though, the GELP system seems to be a significant improvement over head pointing and verbal reporting for giving directional responses.

Without modification, the head-pointing and GELP response methods cannot be used to make distance judgments. Many experiments in audio distance perception have used verbal judgments of distance (Coleman, 1968; Mershon & Bowers, 1979). Studies that have simultaneously examined directional and distance perception have asked subjects to draw the location of the speaker relative to the listener on a sheet of paper (either in azimuth and distance or in both azimuth and distance and elevation and distance) (Butler, Levy, & Neff, 1980). Gilkey suggested using a wire-frame sphere model in the GELP procedure to allow subjects to place the response sensor inside the sphere to indicate distance, but did not test this procedure. For a true three-dimensional localization experiment, drawing responses on paper will be slow and will also require additional time to digitize. Verbal report may be adequate, albeit slow, but no data are available on its accuracy in three dimensions.

Four response methods that were potentially appropriate for collecting three-dimensional locational responses at distances less than one meter were chosen for evaluation in a visual experiment.

1. Direct-Location (DL): In the DL method, a subject simply moves an electromagnetic position sensor directly to the location of the visual target. The location of the target and response can be measured using a spherical coordinate system based on the location of each subject's ears and nose, as described in the Appendix. *A priori*, this appears to be a natural response, since no mental transformation of the target location is required, and subjects can use their own anatomical reference points. Soechting and Flanders (1989) examined the accuracy of pointing to the remembered locations of visual targets with the tip of the finger. They found that when pointing in darkness, estimates of direction were quite accurate but that distance was underestimated. The RMS vector magnitude error from the tip of the finger to the actual target location was 11.7 cm. Surprisingly, performance did not improve when pointing with the fingertip in a lighted room, but did improve substantially when subjects used a pointer instead of the fingertip to indicate the target location. Although the DL response has many desirable properties, no data are available on the accuracy of pointing to targets that

are outside the visual field. It is possible that pointing accuracy will decrease rapidly for sources behind the subject, both because no visual feedback is available and because such locations can be difficult to reach even with a pointer. A different technique that would be roughly equivalent for locations in front and behind the listener was desired.

2. Large-Head Transformation (LH): The LH technique requires the subject to remember the location of the target relative to their own head, and move an electromagnetic position sensor to that same location relative to a full-size Styrofoam manikin head. It was hoped that the anatomical features of this head would allow better judgments of distance and direction than the perfect sphere used for responses in the GELP technique. Because the experiment limited all target locations to the right hemisphere, the head was placed in profile, facing to the right of the subject; thus, response locations in front, to the right, and behind the head were equally accessible. In order to allow responses at distances up to 1 m on the right side, the head was placed 1 m away from the subject. Response locations were recorded using a coordinate system based on the locations of the ears and nose of the manikin head (see Appendix A).

3. Small-Head Transformation (SH): In the SH technique, the full-size Styrofoam manikin head used in the LH technique is replaced with a half-scale soft foam head. When making a response, the subject moves a position sensor to the location relative to the manikin head corresponding to the location of the target relative to their own head, as in the LH response. The major difference between the two responses is that the subjects are expected to scale distance by a factor of one-half in the SH response, allowing the manikin head to be placed closer to the body.

4. Verbal Report (VR): Subjects were familiarized with the spherical coordinate system and asked to give verbal estimates of the azimuth and elevation (in degrees) and distance (in inches) of the target. The VR response is very similar to that used by Wightman and Kistler (1989) in their localization studies.

3.0 EXPERIMENT 1

The first experiment was designed to compare the basic accuracy of the four response methods when there was little uncertainty about the actual target location. A visual target was used, and source locations were limited to the subject's field of vision. The results compare the fundamental accuracy of each of the four response methods.

3.1 Method

3.1.1 Apparatus

Figure 1 shows the overall setup used in the experiment. The experiment was conducted in a large, normally lighted listening booth. The four walls and ceiling of the booth were covered with acoustic foam, and the floor was carpeted. Subjects were seated on a wooden chair near the center of the room and asked to immobilize their heads with the help of a chin-rest constructed from plastic pipe mounted on a heavy base plate. The top of the chin-rest consisted of a lucite block and two plastic screws covered by a soft cloth; the screws provided a reference point allowing the subjects to maintain a consistent head position throughout each block of trials.

A half-scale model of a human head was mounted on the chin-rest in front and below the subject's head. The model was fabricated with soft packing foam and was roughly 10 cm wide at the interaural axis. Although a crude replica, the foam head exhibited most of the basic features of the human head. Eye-sockets and a prominent chin were carved into the solid foam, and a nose, ears and lips were fashioned separately and glued onto the head with rubber cement. A wooden dowel was affixed to the bottom of the head and attached to the chin-rest. This suspended the manikin head 25 cm below and 50 cm in front of the subject's chin. The head was in profile when viewed by the subject, and facing to the subject's right side.

A full-size replica of a human head was also positioned in front of the subject's chair. This manikin head was made of Styrofoam and purchased from a wig shop. The features of

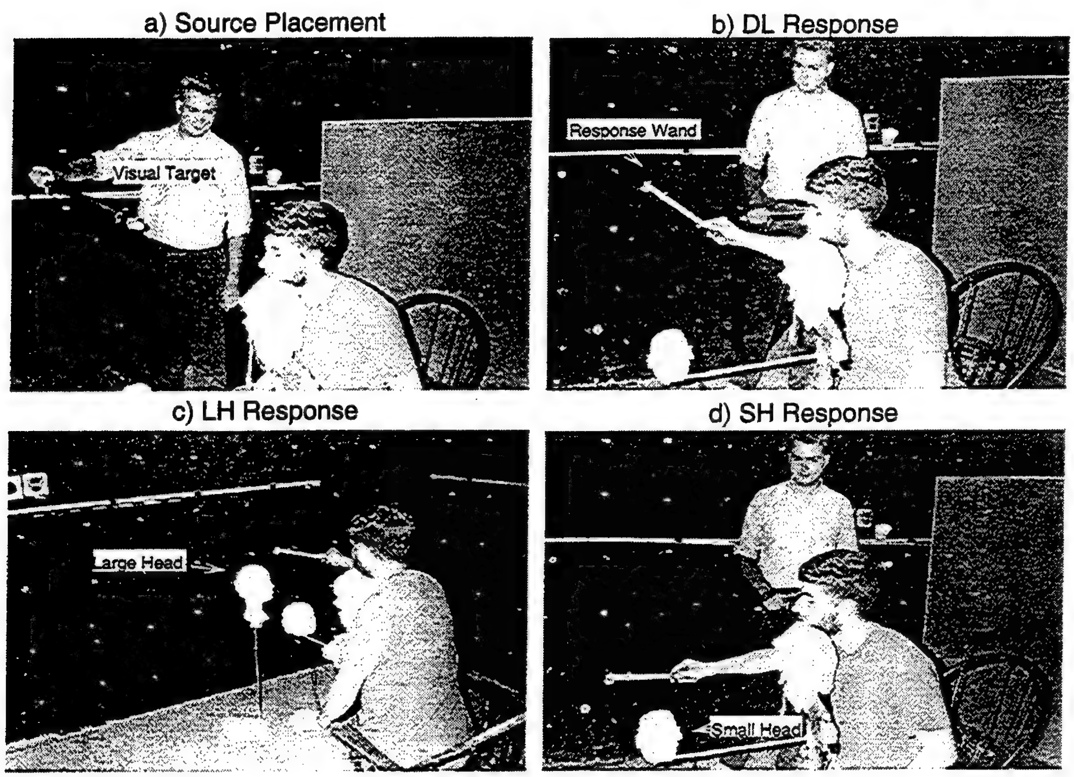

Figure 1: Setup and procedure used for the pointing experiments. The experimenter first places the sensor on the target wand at a random location in the front, right quadrant of the subject, makes sure the subject has seen the location, and presses a response button to have the control computer record that location (a). The target is then moved away and the subject estimates the location of the target by four methods: moving the sensor on the response wand to the location of the target (DL) (b), moving the sensor to the appropriate location relative to the large manikin head (LH) corresponding to the location relative to his or her own head (c), moving the sensor to the location relative to the small manikin head (SH) (d), and verbally reporting the azimuth, elevation, and distance of the source (VR). Note that soft cloth is draped over the chin-rest to enhance the comfort of the subjects.

this head included ears, eyes, nose, mouth, and neck. The full-sized head is relatively small for a human, with an interaural axis length of approximately 16 cm. The head was mounted on a vertical dowel rod in a wooden base, and was placed in profile (facing to the subject's right) 100 cm in front and 25 cm below the subject's chin.

A Polhemus 3-Space Tracker was used to record stimulus and response positions. The source of the Tracker was attached to the chin-rest assembly approximately 20 cm below the subject's chin. One sensor was mounted on the end of a wooden rod, 33 cm long, that was used by the subject to make responses. The other sensor was attached to the end of a clear plastic tube, 58 cm long, which the experimenter used to place the stimuli. The experimenter also used a small hand-held switch to signal when stimulus and response positions should be recorded by a control computer. The Polhemus system is capable of measuring the three Cartesian coordinates of the sensors relative to the source within 0.5 cm in each dimension at distances up to 90 cm, and with slightly degraded accuracy at greater distances.

3.1.2 Calibration

Prior to each block of trials, the subject was asked to find a comfortable position in the chin-rest and immobilize his or her head. Then the Polhemus system was calibrated using nine reference positions. Specifically, the Cartesian coordinates of the left ear, right ear, and tip of the nose were recorded for the subject's head, the large manikin head, and the small manikin head. As discussed in Appendix A, these reference locations were used to establish the coordinate systems used for each of the three heads throughout the block of trials.

3.1.3 Stimulus

The target location in each trial was indicated by the position of a visual pointer. This pointer, a Polhemus head-tracking sensor on the end of a clear plastic wand, was moved by the experimenter to a random position in the front right quadrant of the subject at the beginning of each trial. To determine a random location for the target, the experimenter rolled three fair six-sided dice. One die designated the approximate azimuth of the target (ranging from $0°$ for a one to $90°$ for a six), one die designated the approximate elevation (ranging from $+60°$ for a one to $-60°$ for a six), and one die designated the approximate distance (15 cm for a one to 100 cm for a six).

Table 1: Response order in each block for each subject.

Subject	Block 1	Block 2	Block 3	Block 4
LCL	VLDS	LSVD	DVSL	SDLV
WRD	LSVD	SDLV	VLDS	DVSL
LAR	DVSL	VLDS	SDLV	LSVD
KRJ	SDLV	DVSL	LSVD	VLDS

Note: V=VR; S=SH; L=LH;and D=DL.

Once the target was in position, the experimenter pressed a switch and the computer recorded the coordinates of the target sensor. Two consecutive measurements of target location were made, and the distance between the two measurements was used to ensure that the target was stationary when the switch was pressed. If the target moved more than 1.3 cm between the measurements, a warning tone alerted the experimenter to place the target again. If the target did not move, a different tone alerted the experimenter to move the target away and also informed the subject to begin his or her responses.

3.1.4 Responses

In each trial, the subjects estimated the location of the visual target using four different response methods, DL, LH, SH, and VR, as described in the previous section. In the DL, LH, and SH responses, the subject moved the tip of the response wand to the appropriate location corresponding to the location of the target relative to his or her own head. Once the response sensor was in place, the experimenter pressed the response switch and the coordinates of the response (relative to the manikin head) were recorded. Note that with the half-scale head the subjects scaled distance down by a factor of two, and the Cartesian coordinates of the response (relative to the Small-Head) were doubled to allow direct comparison with the other methods.

The VR response did not require the use of the response wand. The subject simply stated the azimuth, elevation, and distance of the target and these were entered into the control computer by the experimenter.

3.1.5 Subjects and response ordering

Two male and two female volunteer subjects, ranging in age from 21 to 26, participated

in the experiment. Each subject engaged in 4 blocks of 30 trials each, for a total of 120 trials per subject. Each block of trials lasted approximately 45 minutes, and the subjects were allowed 15-minute breaks between blocks. Subjects LCL and LAR completed all four blocks in one day. Subjects KRJ and WRD completed 2 blocks on each of two days.

The response order used in each of the four experimental blocks by each subject is shown in Table 1.

These sequences were selected to minimize the effects of response order on the overall results according to the following criteria:

- The position of each response method was different in every block. Each of the four response methods occurred first in one block, second in one block, third in one block, and fourth in one block.
- No pairs of consecutive response methods were repeated across blocks. The response immediately following a particular response type was different in each of the four blocks.
- The same four response orders were used for each subject's four blocks of trials, but the order of the blocks was different for each subject.

3.1.6 Raw data

Figures 2 to 4 show the stimulus-response pairs in azimuth (Figure 2), elevation (Figure 3), and log distance (Figure 4), respectively, for each subject and each response method. In each panel, "correct" (i.e., veridical) responses are indicated by a solid line. From the raw data shown in the left panels of these figures, three observations can be made. First, there are substantial differences between the accuracy of the response methods in each dimension. In azimuth, for example (Figure 2), the spread of responses is much greater for the LH method than for any other response method. Second, biases are evident in that the responses are often clustered away from the correct response, particularly in the LH, SH, and VR methods. For example, in distance with the VR method (Figure 4), the responses cluster below the solid line, indicating a bias to underestimate the target distance. Third, and of greatest significance, the DL method has the least response variability and the least bias of the four methods, with responses clustered closely around the correct response in azimuth, elevation, and distance.

3.2 Results

Five measures of response accuracy were used to quantify the results from the four response methods. These measures (Table 2) summarize the mean errors and standard deviations for the following error parameters:

- The signed error in *azimuth*: The difference between the azimuth location of the response and the azimuth location of the target. Recall that $-90°$ is directly right of the subject and all targets were on the right side, so a negative value for this error parameter implies the response was more lateral than the target.
- The signed error in *elevation*: The difference between the elevation location of the response and the elevation location of the target. Note that $90°$ is directly above the subject, so a negative value of signed elevation error implies the response is below the target.
- The signed percentage *distance* error: The percent difference between the target distance and the response distance $((\frac{r_{response}}{r_{target}} - 1) \cdot 100\%)$. A negative value indicates the subjects have underestimated distance.
- The *angle* error: The magnitude of the angle of the arc between the target and response locations on a great circle centered at the origin. Note that this error depends on both the stimulus azimuth and elevation and the response azimuth and elevation.
- The *vector-length* error: The ratio of the length of the vector going from the location of the response to the location of the target divided by the distance from the center of the head to the target, expressed as a percentage. The vector-length error includes both directional and distance components and measures overall performance in the task.

These five errors can be divided into two categories. The signed azimuth, elevation, and percentage distance errors can be positive or negative according to the direction of the error. The mean value of these quantities represents the bias of the responses, while the standard deviation measures consistency. In contrast, the angle and vector-length errors are strictly positive error measurements. The mean values of these parameters represent the total error, including any response bias and the spread of the responses around the mean, while the standard deviations are useful primarily for evaluating the significance of changes in the

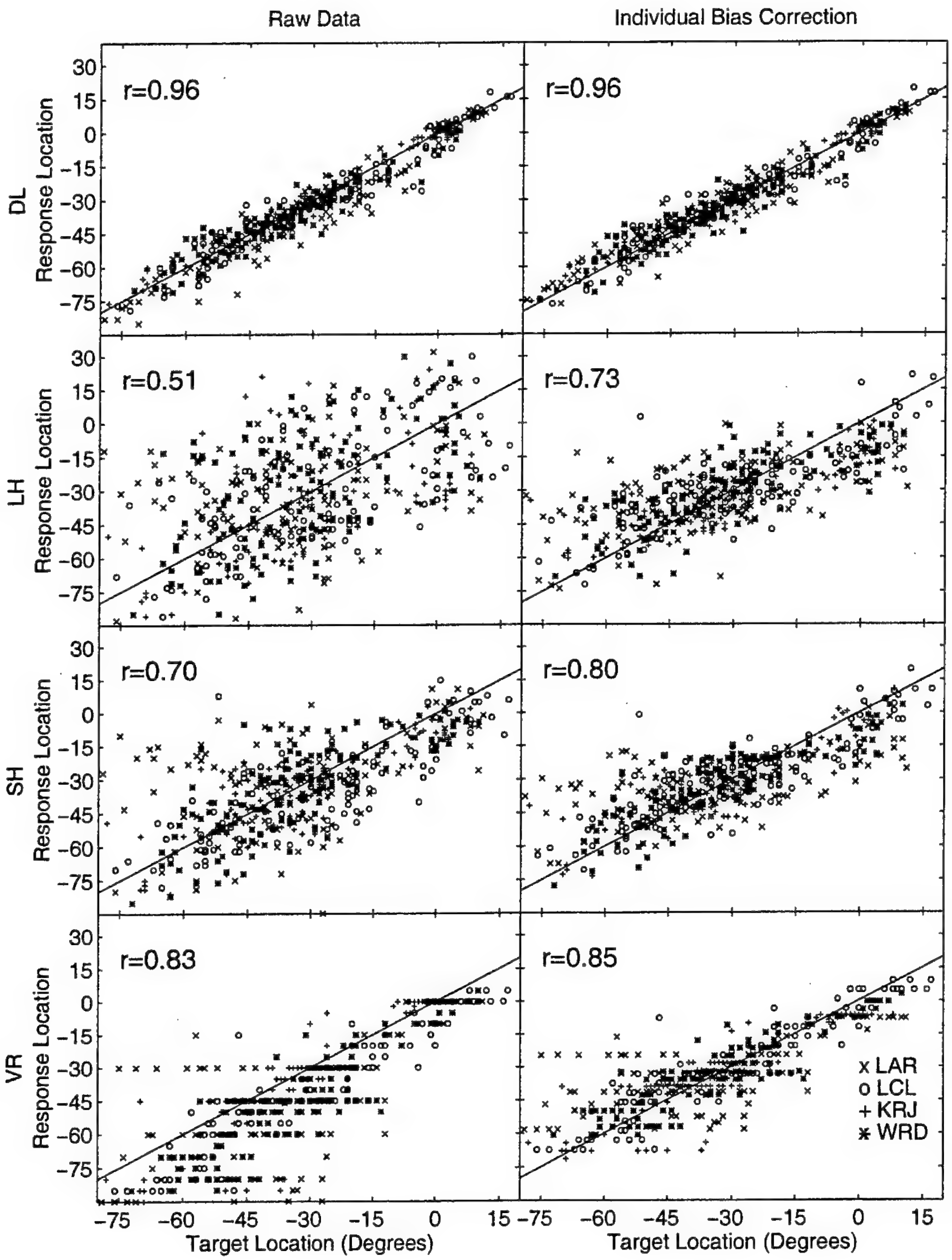

Figure 2: Stimulus-response pairs for azimuth. Each row shows results for a different response method (as labeled). Raw data are plotted in the left panels; data corrected for individual response biases are plotted in the right panels (see text for details). Different symbols are used to represent the responses of each subject. The solid line indicates "correct" responses. Note that the responses for verbal report are quantized (along the ordinate), representing subject biases in favor of particular response values. The azimuths range from −90° directly to the right of the subject to 0° directly in front of the subject.

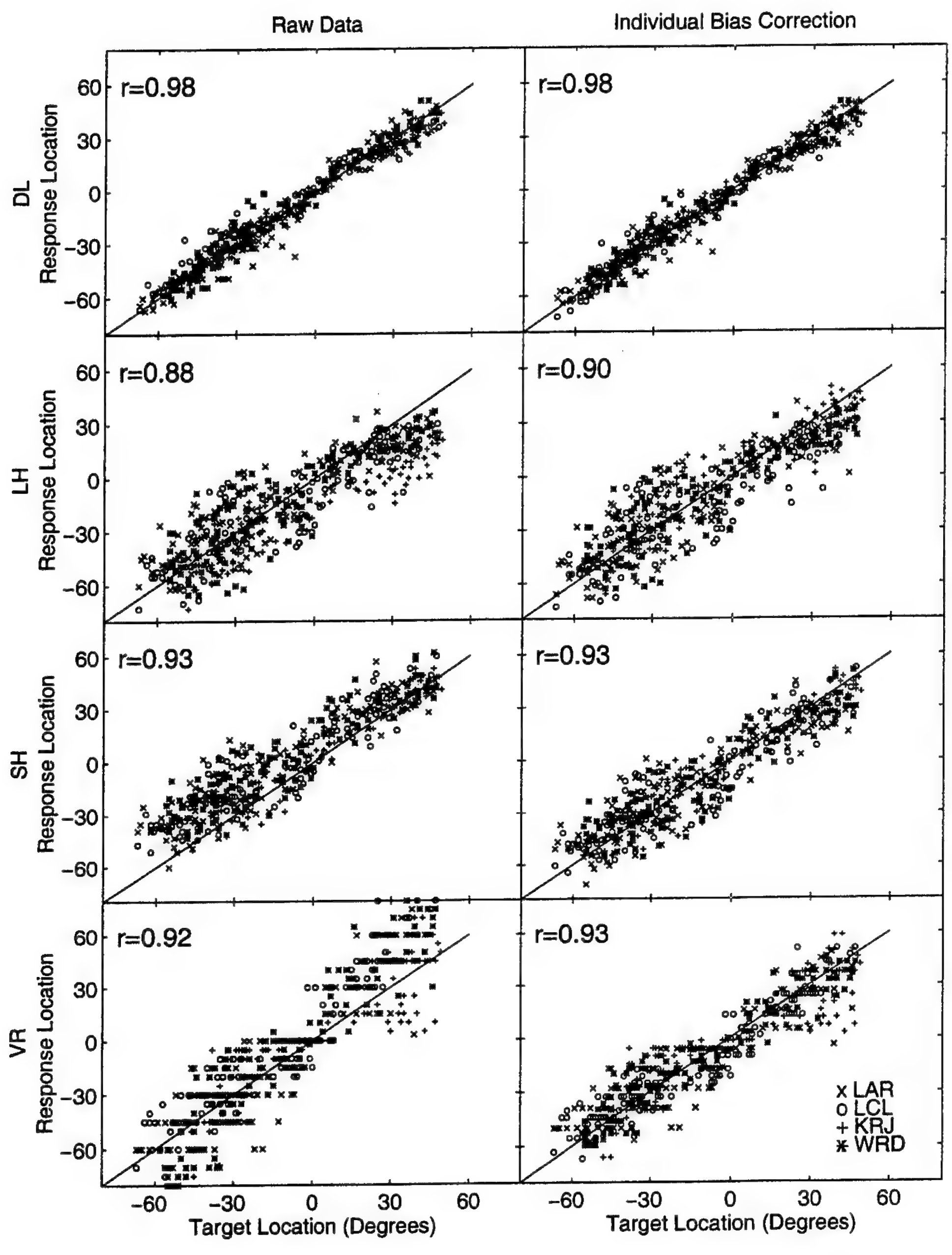

Figure 3: Stimulus-response pairs for elevation. Other details as in Figure 2. Elevations range from −90° directly below the subject to 90° directly above the subject.

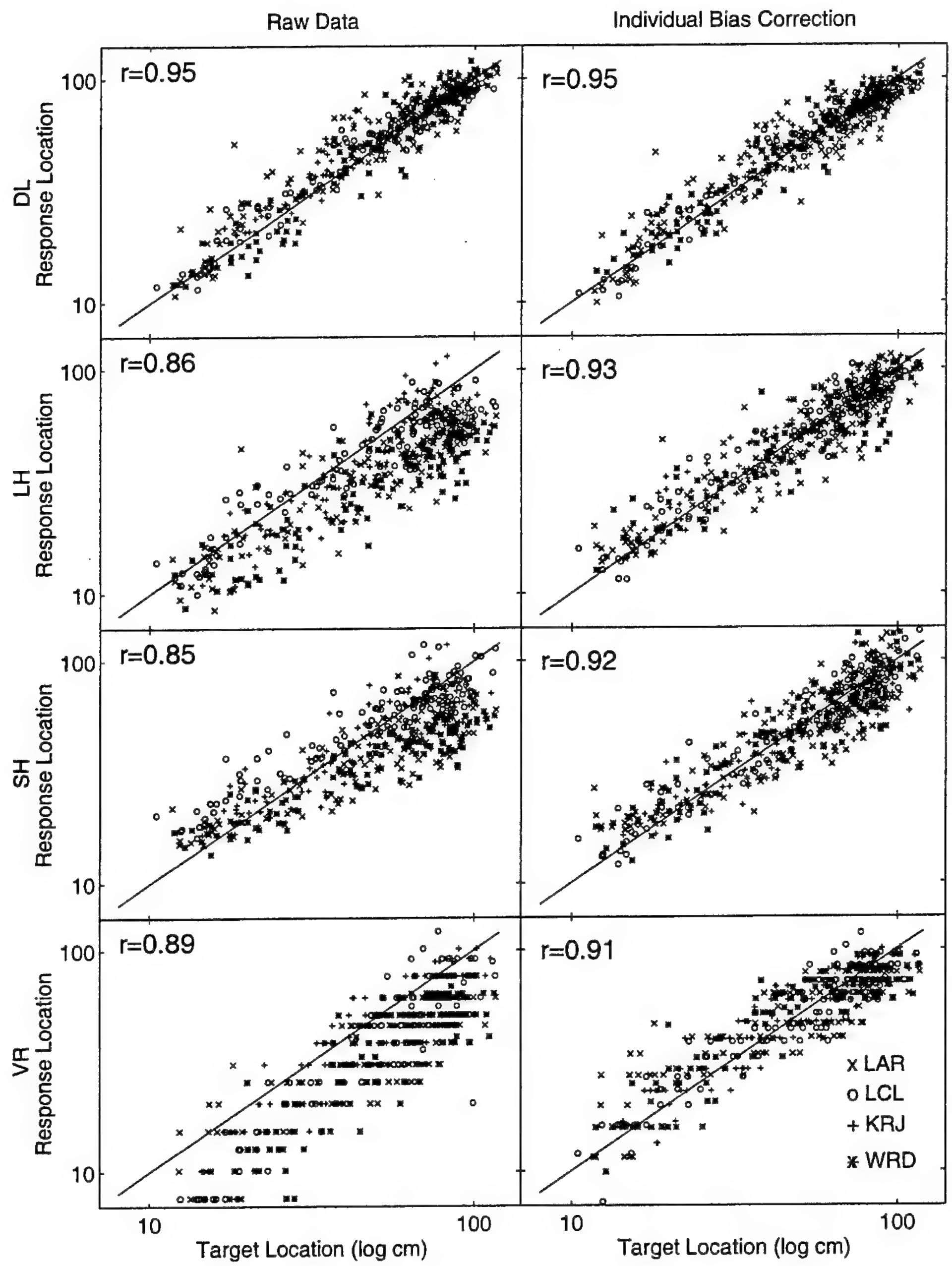

Figure 4: Stimulus-response pairs for distance. Other details as in Figure 2 but for distance rather than azimuth. Distances range from 10 cm to 100 cm from the center of the head.

Table 2: Mean errors for each response method

Error Type	DL		LH		SH		VR	
Azimuth	$-1.8°$	(5.2)	$0.5°$	(22.1)	$-1.8°$	(15.1)	$-9.6°$	(12.6)
Elevation	$0.5°$	(5.8)	$-4.9°$	(13.9)	$10.6°$	(11.6)	$4.7°$	(15.9)
Distance	7.6%	(19.9)	-21.8%	(23.7)	-9.6%	(26.1)	-26.2%	(22.4)
Angle	$6.0°$	(4.2)	$21.7°$	(9.5)	$17.6°$	(9.4)	$18.7°$	(10.6)
Vector-Length	21%	(16)	48%	(15)	41%	(18)	45%	(16)

Note: Standard deviations are in parentheses.

mean value. When responses are unbiased, or when bias is removed as described in the next section, the mean angle and vector-length errors measure only the spread of responses and become similar to the standard deviations of the signed errors.

A note should be made about the interactions between the azimuth, elevation, and angle errors. The azimuth error is potentially flawed, because it increases in sensitivity as the target moves away from the horizontal plane. The interaction can be visualized by imagining the spherical coordinates of a world traveler. The angle, or great-circle, measure of the traveler's movements is directly proportional to the absolute number of miles he or she has moved from the starting point, and does not depend on the starting location. Similarly, the change in elevation is proportional to the number of miles north or south the traveler moves and is independent of starting location. The change in azimuth is trickier; it is proportional to the number of miles between the points on the equator due south (or north) of the traveler's starting and ending positions. The change in azimuth generated by moving one mile due west (or east) is inversely proportional to the cosine of the elevation, so the change in azimuth caused by moving one mile west near the north pole is much greater than the change in azimuth caused by an identical movement near the equator. As a result, azimuth accuracy may decrease as elevation magnitude increases. For this reason, the angle error is a better measure of accuracy than the azimuth error in spherical coordinates. Also, note that errors in elevation that cause the response to move across the top of the unit sphere will cause both the azimuth and elevation errors to be meaningless, but will not affect the angle error (although the range of elevations in this experiment is limited to $-60°$ to $60°$ so polar transversals are not a problem here).

As can be seen in Table 2, the azimuth response is essentially unbiased except in the VR paradigm, where an average bias of almost $10°$ toward the right side occurs. For elevation, the largest bias is in the SH paradigm, where subjects on average responded more than $10°$

above the true target location. The other biases in elevation were less than 5° in magnitude. In both azimuth and elevation, DL has less than 2° of bias. The signed distance errors show that, on average, subjects overestimated distance in the DL response, but underestimated distances when using the other response methods.

For angle error, an ANOVA on the raw data shows that the main effect of response method is highly significant ($F_{3,1650} = 425.12$, $\alpha = 0.0001$). Pairwise t-tests (at the $\alpha = 0.01$ level) show that DL is the best response angle response method (mean error=6.0°), SH and VR are not significantly different (mean errors=17.6° and 18.7°), and LH is the worst method (mean error = 21.7°).

As the final measure of accuracy, the vector-length error shows again that DL is superior to the other pointing methods (mean error = 21%), followed by SH (mean error=41%), then VR (mean error 45%) and finally LH (mean error = 48%). The DL method is clearly best, and the differences between the other three methods, though small, are also significant (one tailed t-test, α=0.01).

The overall results can be summarized as follows:

- The DL response method is greatly superior to all other methods, and is essentially unbiased in direction.
- The LH, SH, and VR responses all exhibit some type of directional bias (in azimuth, elevation, or both) and all underestimate distance.
- The overall angle errors for LH, SH, and VR are all large, although LH is slightly (but significantly) worse than the other two methods.

3.2.1 Inter-subject variability (raw data)

Although there were some differences in performance among the four subjects, they performed quite similarly in general (see Figure 5). For both the angle (great circle) error and the vector-length error (the top two panels of Figure 5), all four subjects exhibited the best performance (smallest mean error) in the DL method. For the other response methods, three of the four subjects performed similarly. The fourth subject (KRJ) showed significantly better performance with the VR and SH methods than with the LH method.

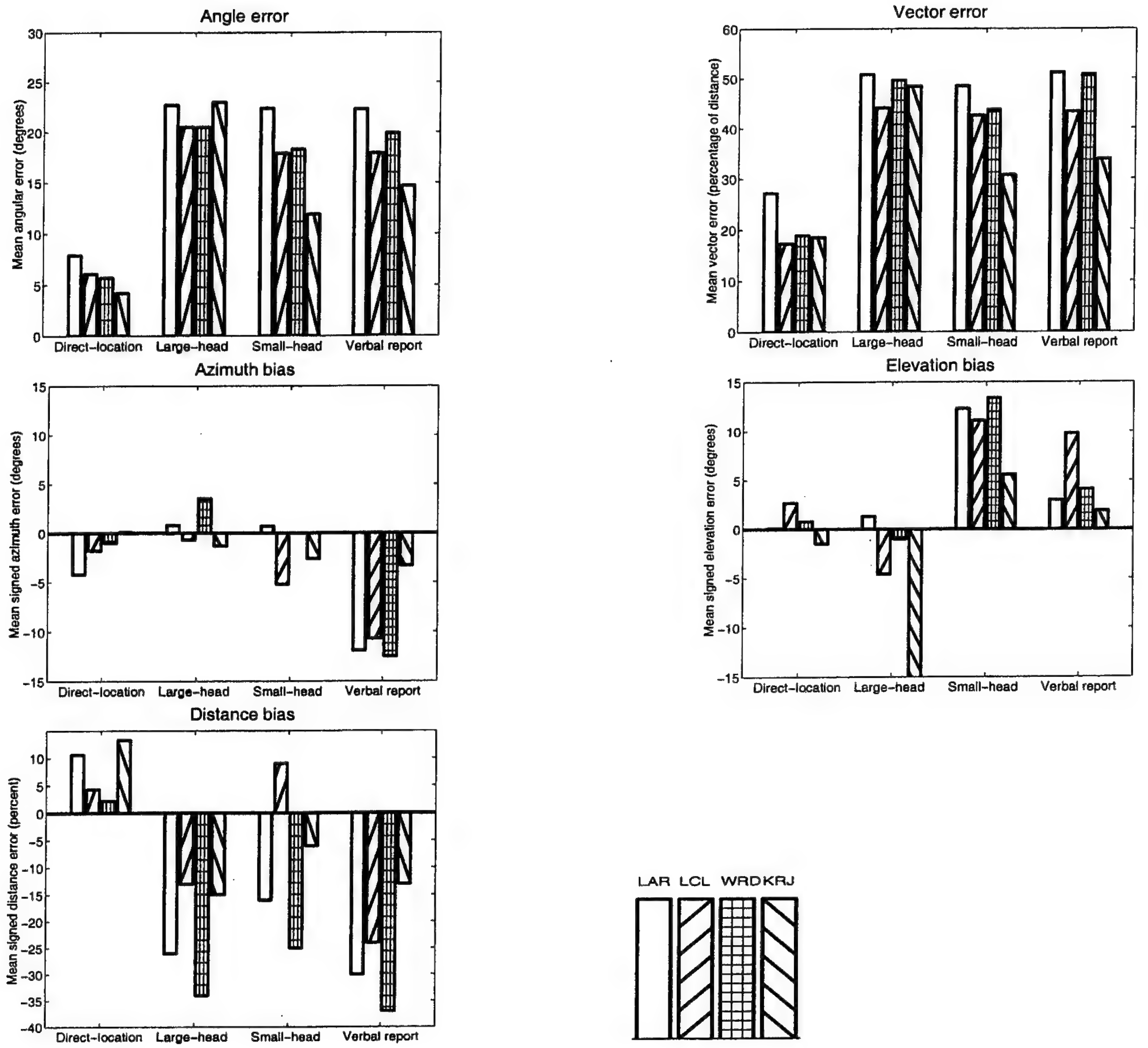

Figure 5: Mean errors for each subject and each response method

The biases in each response method were also generally similar among the subjects (the bottom three panels of Figure 5). In every response method and every response dimension, at least three of the four subjects exhibit the same general bias trend (negligible bias, positive bias, or negative bias). In three cases, subject KRJ deviates from the general bias trend. KRJ has a much smaller bias in azimuth than the other subjects in the VR method, and his mean elevation response was significantly below that of the other subjects in the LH and SH methods. Since the other subjects exhibit small elevation biases in the LH response and large positive elevation biases in the SH response, KRJ's unique trends make him the least biased subject in the SH method and the most biased subject in the LH method. These differences in bias help explain why KRJ's mean angle and vector-length errors (which include bias effects) are smaller in the VR and SH methods than in the LH method.

A difference in head positioning during the experiment may explain the unusual elevation response biases exhibited by subject KRJ. An examination of the orientation of the subjects' heads during calibration indicates that KRJ's head was tilted slightly down (i.e., his nose was slightly lower than his interaural axis), while the other subjects' heads were tilted slightly up. The effect of this tilt is clearly seen in Table 3, which shows the mean stimulus target locations for the four subjects. While the average target azimuths and distances were similar for the three subjects, the average target elevation for KRJ was 13.3° higher than the average target elevation for the other three subjects. This difference suggests that KRJ's head was tilted down slightly during the experiment, and that the experimenter did not correct for this tilt when positioning the visual target. If a subject's internal representation of the target location did not account for the tilt of his head, he would position the target sensor at a lower elevation relative to the manikin head when his head was tilted downward than when his head was tilted upward, while the response bias in the DL method would be unchanged. The bias trends in Figure 5 are consistent, therefore, with KRJ's head having been tilted downward and the other subjects' heads tilted upward, although it is not clear why the VR response is apparently unaffected.

3.3 Bias Correction

3.3.1 Overall correction

In every response method except DL, the subjects generally performed quite poorly. Angle errors were close to 20°, and vector-length errors were approximately 45%. In part,

Table 3: Mean location of the visual target for the stimuli used with rach subject

Subject	Azimuth	Elevation	Distance
LAR	$-33.8°$	$-15.2°$	54.5 cm
LCL	$-29.3°$	$-10.8°$	57.8 cm
WRD	$-32.4°$	$-12.1°$	57.4 cm
KRJ	$-31.8°$	$0.6°$	56.2 cm

Note: Values are given based on the coordinate system determined for each subject. With the exception of elevation of subject KRJ, the mean locations are roughly constant across subjects.

Table 4: Bias correction parameters and correlation coefficients

	DL			LH			SH			VR		
	m	b	r_{corr}	m	b	r_{corr}	m	b	r_{corr}	m	b	r_{corr}
Azimuth	0.97	0.9	0.96	0.42	-18.6	0.51	0.71	-8.6	0.70	0.69	-3.2	0.83
Elevation	1.00	-0.5	0.98	1.05	5.5	0.88	1.05	-10.6	0.93	0.73	-5.9	0.92
Distance	0.98	0.0	0.95	0.93	0.5	0.86	1.06	-0.1	0.85	0.70	1.1	0.89

Note: m gives the slope and b gives the y-intercept of the linear bias corrections. The transformed value of response x in each case is $mx + b$. The correlation (r_{corr}) between target location and response location is also shown.

these overall errors were caused by biases in the responses. In VR, for example, responses were on average approximately 10° too far to the right, and 26% too close (Table 2). Each of these biases contributed to the average errors of 17.6° for angle and 45% for vector-length. If biases were systematic, the accuracy of a given response method could be enhanced by transforming the biased response locations into non-biased estimates of actual source location.

A linear transformation of each response coordinate was used to compensate for response biases. The coefficients of these linear transformations were found from the linear regression of the actual target coordinates (independent variable) on the response coordinates (dependent variable), separately for azimuth and elevation. The distance transformation was based on the regression of the log of the actual distance on the log of the response distance. The parameters of the overall bias corrections are shown in Table 4.

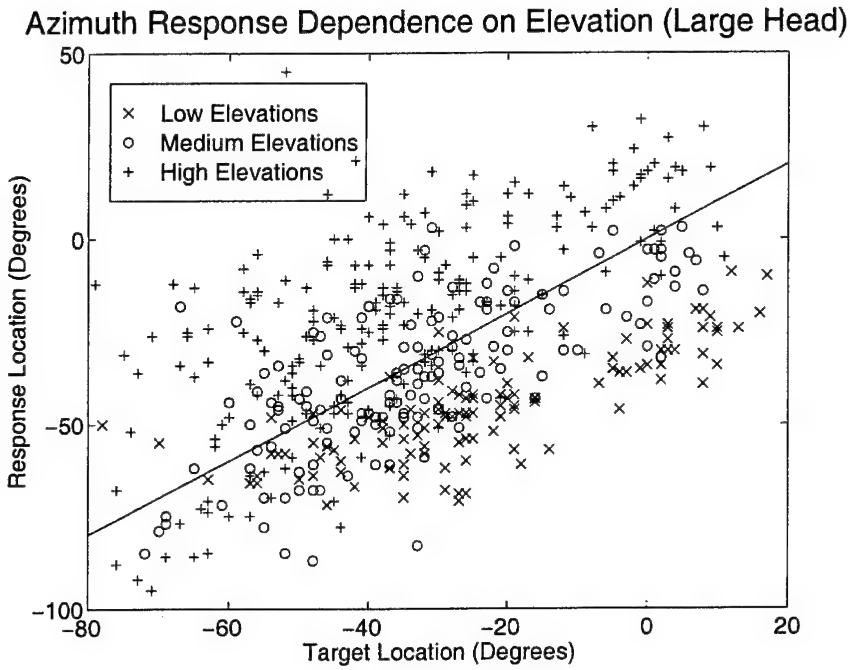

Figure 6: Stimulus and response azimuth for the LH method, with elevation as a parameter. The three symbols represent bias-corrected elevation responses less than $-15°$ (low), between $-15°$ and $15°$ (medium), and greater than $15°$ (high). Note that the azimuth response has a negative bias at low elevations (with x's falling below the zero-bias diagonal), is approximately unbiased at medium elevations, and is positively biased at high elevations.

Table 5: Elevation-dependent bias correction parameters for azimuth

	LH			SH		
Elevation	m	b	r_{corr}	m	b	r_{corr}
$\leq -15°$	1.11	26.6	0.80	0.86	6.1	0.83
$\leq 15°$	0.70	-6.3	0.75	0.81	-2.3	0.82
$> 15°$	0.51	-25.7	0.65	0.66	-15.6	0.69
Corrected			0.73			0.77

Note: Elevations were corrected for linear bias before sorting into bins. See text for details.

3.3.2 Elevation dependent correction

The bias correction can be improved further by exploiting interactions that occur across response dimensions. An analysis of the biases in azimuth, elevation, and distance revealed that response biases for elevation were essentially independent of both distance and azimuth for all four response methods. The response biases in azimuth were independent of distance, but were strongly dependent on elevation for the LH response (see Figure 6) and slightly dependent on elevation for the SH responses. Similarly, the response biases for distance were dependent on elevation for the LH and SH responses. Since the elevation responses are independent of target azimuth and distance, it is possible to first estimate the target elevation by transforming the elevation response, and then transform the azimuth and distance responses based on this unbiased estimate of elevation. For simplicity, the estimate of elevation was used to divide the response region into three bins — one for corrected elevations below $-15°$, a second for elevations from $-15°$ to $15°$, and a third for elevations above $15°$. Linear regressions of target azimuth on response azimuth were performed in each of the three bins, and the resulting coefficients were used for bias correction in azimuth (Table 5). A similar procedure was used to determine the bias correction coefficients for the log of the response distance (Table 6). The elevation-dependent bias correction causes the correlation of stimulus and response azimuth to increase from 0.51 to 0.73 for the LH method and from 0.70 to 0.77 for the SH method (see Tables 4 and 5), and slightly increases the stimulus-response correlation in distance for each method.

3.3.3 Subject dependent correction

When the individual subject data were examined it was found that one subject sometimes deviated from the overall trend (see Figure 5). It is possible to correct the responses for these

Table 6: Elevation-dependent correction parameters for log distance

	LH			SH		
Elevation	m	b	r_{corr}	m	b	r_{corr}
$\leq -15°$	0.96	0.2	0.92	0.96	0.1	0.85
$\leq 15°$	0.89	0.6	0.89	1.12	-0.2	0.89
$> 15°$	0.93	0.7	0.89	1.06	0.1	0.84
Corrected			0.91			0.88

Note: Elevations were corrected for linear bias before sorting into bins. See text for details.

individual subject biases. This was done in the same way as the elevation-based corrections described in the previous section, except that different correction parameters were calculated for each individual subject. This results, of course, in a large set of correction coefficients, and these coefficients will not be shown. The individually corrected data were used to calculate vector-length and angle errors for comparison with the non-individualized bias-corrected results. The correspondence between the subject responses when bias corrected for individual differences and the actual target locations is shown in the right panels of Figures 2, 3, and 4. In all cases the corrected results are clearly superior to the raw data shown in the left panels, with the responses more tightly clustered around the diagonal representing correct responses.

3.3.4 Bias correction summary

The angle and vector-length errors were calculated with each level of bias correction (Table 7). The errors decrease to varying degrees as each additional level of complexity is added to the bias correction scheme:

- The DL response method changes insignificantly with bias correction, indicating that these responses were essentially unbiased without any corrections.
- The LH, SH, and VR responses are all improved by bias correction, but remain roughly comparable in overall performance.
- Even with bias correction, the accuracy of the LH, SH, and VR methods remains substantially inferior to the DL method.

Table 7: Response errors for different levels of bias correction

Error	Method	No Correction	Overall	Overall + Elevation	Individual + Elevation
Angle	DL	6.0°	6.0°		5.6°
	LH	21.7°	18.6°	16.5°	15.4°
	SH	17.6°	14.6°	13.6°	13.0°
	VR	18.7°	13.2°		12.4°
Vector-Length	DL	20%	19%		19%
	LH	48%	46%	39%	35%
	SH	41%	41%	37%	33%
	VR	45%	35%		33%

- The majority of the performance improvement is obtained with the overall and elevation-based bias correction. This joint correction captures more than 80% of the total improvement in angle error, and at least 50% of the total overall improvement in vector-length error. The additional improvement provided by correction for individual subject biases is small.

These results indicate that a linear bias correction can improve the accuracy of the indirect response methods. Furthermore, it appears that a general bias correction scheme chosen for all subjects will work almost as well as a more complex scheme based on the individual peculiarities of each subject's responses.

3.4 Locational dependence of response errors

A priori, one might expect the accuracy of the response to depend on the location of the visual target. For example, the angle error might be larger for targets off to the side of the subject than for targets directly in front of the subject. In order to investigate this possibility, the correlation coefficients for the response error and the azimuth, elevation, and distance of the stimulus location were computed. The data were first corrected for bias using the overall elevation-adjusted bias correction discussed previously. The results (Table 8) are the correlation coefficients between the error variable (angle and vector-length) and the azimuth, elevation, and distance of the target. The angle errors are essentially uncorrelated with target position: the largest correlation coefficient for angle has magnitude 0.17, and the average magnitude of the correlation coefficient is 0.09. The vector-length errors are

Table 8: Correlation coefficients between response errors and target location

Error	Variable	DL	LH	SH	VR
Angle					
	Distance	-0.13	-0.11	0.03	-0.16
	Elevation	0.02	0.06	-0.14	0.08
	Azimuth	-0.17	-0.08	-0.10	-0.03
Vector-Length					
	Distance	-0.27	-0.34	-0.19	-0.32
	Elevation	0.03	0.02	-0.08	0.06
	Azimuth	-0.11	-0.10	-0.09	0.09

also essentially uncorrelated with elevation and azimuth ($|r_{corr}| \leq 0.11$, mean magnitude 0.07), but negatively correlated with distance for all four response methods (average -0.28). This indicates that the vector length error is not simply proportional to distance and may include a constant term that contributes more to the error ratio at close distances than at far distances. Overall, however, it appears that the specific location of the target within the visual field has little influence on performance.

3.5 Memory Effects

On each trial, subjects were required to use all four response methods for each stimulus. While this ensures that each response method used exactly the same stimulus set, it presents the possibility that the accuracy of a response could depend on its position in the response sequence. *A priori*, one expects the subject's mental image of the target location to degrade over time; hence, the first response given after each stimulus might be the most accurate, and the last response the least accurate. In our experiment, four different response orders were used, and each response method was first, second, third, and fourth in an equal number of trials.

As a simple assessment of response order effects, we compared the accuracy of each response method in trials when it was the first response and trials when it was the last response (Table 9). The DL response degrades the most when it follows the other response methods. The average DL angle error increased 80% and the vector-length error increased 130% when DL was the last response versus the first response. For LH, the average angle and vector errors also increased significantly, but not as dramatically as in the DL response. None

Table 9: Accuracy of first and last responses

Error	Method	First Response		Last Response	
Angle					
	DL†	4.0°	(3.0)	7.2°	(4.8)
	LH†	19.8°	(9.5)	24.8°	(9.4)
	SH	17.0°	(9.2)	17.2°	(9.8)
	VR	19.0°	(10.9)	19.3°	(10.9)
Vector-Length					
	DL†	12%	(7)	28%	(20)
	LH†	45%	(14)	51%	(15)
	SH	41%	(15)	42%	(19)
	VR	47%	(15)	46%	(15)

Note: The data were not corrected for biases. Standard deviations are given in parentheses. The † indicates differences significant at the $\alpha = 0.001$ level (two-tailed t test); no other differences are significant at the $\alpha = 0.05$ level.

of the other errors increased significantly. A likely explanation for these results is that the DL response relies on explicit memory of the source's visual location, and this memory degrades when one is distracted by making other intervening responses. In contrast, the three indirect response methods require one to encode the location of the source and mentally transform this location into either verbal coordinates or coordinates relative to the manikin heads. This encoded memory may be less volatile than the explicit visual memory. By analogy to the memory model developed by Durlach and Braida (1969), the memory model used for DL is a *trace memory* that degrades over time and that used for the three indirect responses is a *context coded memory* that is temporally invariant. Suppose the same encoded version of the source location was used for all three indirect response methods. Then one would expect the same types of errors in all three methods within a given trial. This hypothesis was tested by examining the correlations between the errors of the three indirect response methods in each trial. Correlation coefficients between the LH and SH response errors were found to be very high – 0.79 for azimuth errors, 0.78 for elevation errors, and 0.76 for log-distance errors. Correlation coefficients between the errors of the other response methods did not exceed 0.47 (and the average of the other correlations was only 0.31). It is therefore likely that the same mental encoding of location was used for both the LH and SH responses (not surprisingly, considering the similarities between the tasks), but this encoding does not appear to be the same as that used for the VR responses.

4.0 EXPERIMENT 2

Overall, none of the three indirect response methods was found to be particularly attractive. The accuracy of the LH, SH, and VR responses was similar when (subject-independent) bias correction was applied. Other factors do differentiate the methods, however. The VR method was found to be too slow to be practical (often taking as long as the other three responses combined), and the LH response showed marginally larger angle errors than the other indirect methods. Consequently the SH response method was tentatively chosen as the best of the indirect methods. Since the small manikin allows subjects to scale distances down by a factor of two, it is not necessary to place the manikin in profile (as was done in Experiment 1) in order to allow the subjects to reach all locations in the right hemisphere at distances up to 1 m. That is, the head could be placed facing in the same direction as the subject. We hypothesized that this might yield improved performance by simplifying the internal transformations made by the subject, since a 90° reference frame rotation is eliminated. A simple experiment was designed to test this hypothesis.

4.1 Method

The second experiment was similar to the first experiment with only a few exceptions. First, testing was performed in an open laboratory space and not a small listening booth. Second, the responses were restricted to the SH technique. The small manikin head was no longer attached to the chin rest, but was mounted on the vertical stand previously used for the large manikin head. This allowed the Small-Head to be rotated to face either the same direction as the subject or directly to the right of the subject (i.e. in profile as in Experiment 1).

Four subjects (three male, one female, ranging in age from 23-30) were paid to participate in the experiment. Four blocks of 50 trials were collected from each of the subjects. For two of the subjects, the head was in profile for the first and third blocks and facing forward in the second and fourth blocks. For the other two subjects, the opposite ordering was used. Each block took about 20 minutes, and subjects were given a break between blocks.

Table 10: SH Results from experiment 1 and experiment 2

Error Type	Experiment 1		Experiment 2	
Azimuth	$-0.8°$	(12.1)	$-3.4°$	(12.7)
Elevation†	12.5°	(10.5)	9.7°	(9.1)
Distance††	-16.5%	(23.5)	-0.6%	(21.2)
Angle	16.9°	(9.1)	15.6°	(7.7)
Vector-Length††	41%	(15)	35%	(15)

Note: Comparison of the average response errors for the SH response in Experiment 1, when it was the first response in a trial, to the response errors for the profiled head condition of Experiment 2. Standard deviations are in parentheses. The † and †† indicate differences significant at the $\alpha = 0.05$ and $\alpha = 0.001$ level respectively (two-tailed t test).

4.2 Comparison to Experiment 1

The data collected with the head in profile from Experiment 2 were first compared to the results from Experiment 1 for the trials in which the SH response was the initial response (Table 10). The similarity between the angle biases and the overall angle error (and their associated standard deviations) is striking. In each experiment, the subjects showed a small negative bias in azimuth and a positive bias of approximately 10° in elevation. In both cases, the overall angle error was near 16°. Only the difference in elevation bias was statistically significant ($\alpha = 0.05$, using a two-tailed t test).

In contrast, the bias for distance and the mean vector-length errors were both significantly smaller in Experiment 2 than in Experiment 1. This could be a result of reduced response complexity since subjects only needed to think about one response, or from variations among subjects.

The availability of SH response data from a second group of subjects allows us to compare the best overall bias correction of the two groups. As before, this correction was calculated from the linear regression of the target location on the response location. The results (see Table 11) are very similar for both experiments. Although the slope of the bias correction in elevation was slightly higher in the first experiment, the corrections are otherwise nearly identical. Interestingly, the optimal bias correction in distance is identical even though distance accuracy was significantly better in Experiment 2.

The individual subject biases (not shown) were not as consistent as the average biases for the two groups. Individual biases in azimuth for the SH response (in both experiments)

Table 11: Bias correction parameters and correlation coefficients in experiments 1 and 2.

	Experiment 1			Experiment 2		
	r_{corr}	m	b	m	b	r_{corr}
Azimuth	0.71	-8.6	0.70	0.72	-6.8	0.75
Elevation	1.05	-10.6	0.93	0.92	-9.3	0.95
Distance	1.06	-0.1	0.85	1.06	-0.1	0.92

Note: The transformed value of response x in each case is $mx + b$, where m is the slope and b is the y-intercept. The correlation (r_{corr}) between target and response locations is also shown.

Table 12: Mean responses for forward-facing and profiled SH

Error Type	Forward-Facing		Profiled	
Azimuth††	0.6°	(11.6)	−3.4°	(12.7)
Elevation††	7.0°	(9.7)	9.7°	(9.1)
Distance	-1.8%	(26.4)	-0.6%	(21.2)
Angle	15.9°	(7.9)	15.6°	(7.7)
Vector-Length†	38%	(18)	35%	(15)

Note: Standard deviations are in parentheses. The † and †† indicate differences significant at the $\alpha = 0.05$ and 0.001 levels, respectively (two-tailed t test).

ranged from −10.6° to 2.7°, and the biases in elevation ranged from 5.6° to 19.7°. Standard deviations for both azimuth bias and elevation bias were 4.9°. Despite these individual differences, the mean results from the two experiments support the idea that the SH response can be improved substantially with a subject-independent bias correction scheme.

4.3 Comparison of Profiled and Forward-Facing SH

The principal hypothesis motivating Experiment 2 was that the forward-facing condition would yield better performance than the profiled condition. The results clearly reject this hypothesis (see Table 12). In fact, the profiled head condition was found to be slightly, but significantly, more accurate in vector-length error. Overall, however, the difference between the two manikin orientations appears negligible, and this difference diminishes further when the data are corrected for overall biases.

5.0 EXPERIMENT 3

In Experiments 1 and 2, the DL response was by far the best response method. The SH response, which was suggested as the best alternative to DL, produced angle and vector-length errors 2-3 times as large as the DL response even when corrected for bias. The DL method was so much better than the indirect response methods in the front hemisphere that it seemed possible that it might also be superior in the rear hemisphere, despite the difficulties of reaching behind the body and outside the visual field.

In order to test response accuracy for locations outside the visual field, a different (non-visual) stimulus was necessary. Since the primary intent of these experiments was to evaluate response measures appropriate for a near-field psychoacoustic experiment, an auditory stimulus was chosen. While such a stimulus was expected to generate a noisier response estimate than a visual target, it would allow useful comparisons of the accuracy of DL inside and outside the visual field.

5.1 Method

There were two major differences between Experiment 3 and Experiments 1 and 2. First, the stimulus range was expanded from the front right quadrant to the entire right hemisphere. An approximate azimuth location of the stimulus was still generated by rolling a six-sided die on each trial, but the range was increased to $0°$ to $-180°$. Second, in Experiment 3 an acoustic point source was used to indicate the target location. Subjects were asked to keep their eyes closed while the source was being placed prior to each trial. Then a 125-ms burst of white noise was generated at the random target location. The point source was moved out of the way, and the subjects opened their eyes and gave estimates of the source location with the DL and SH responses. The subjects used the chin-rest to keep their heads stationary throughout the experiment.

Two subjects were tested, one male and one female, both of whom had previously participated in Experiment 1. Four blocks of 50 trials were collected from each subject, but a few trials had to be discarded because of head-tracker failures, so a total of 392 trials are usable.

Table 13: Comparison of response accuracy in the front and rear hemispheres for the DL and SH responses in experiment 3

	DL				SH			
Error	Front		Rear		Front		Rear	
Azimuth	$-4.9°$	(10.7)	$-0.1°$	(11.1)	$-5.7°$	(17.1)	$-14.8°$	(13.6)
Elevation	0.2°	(11.4)	4.4°	(11.5)	8.3°	(13.2)	1.5°	(13.0)
Distance	11.9%	(32.5)	8.9%	(29.7)	-1.5 %	(33.8)	1.6%	(34.4)
Angle	11.9°	(8.4)	13.9°	(8.0)	16.9°	(11.7)	20.7°	(9.6)
Vector-Length	37%	(24)	38%	(21)	43%	(23)	48%	(24)

Note: The stimulus was the location of a short sound signal, a 125-ms burst of white noise.

Each block of trials took approximately 30 minutes. In some trials subjects misperceived sound sources in the front hemisphere as arising from the rear hemisphere, a phenomenon that has been reported frequently by others. Whenever both the DL and SH response methods indicated a reversal had occurred, the responses were 'corrected' by reflecting the subject response across the front-back plane, as discussed by Wightman and Kistler (1989).

5.2 Results and Discussion

5.2.1 Analysis of psychoacoustic experiment

Although the primary goal of Experiment 3 was to verify the validity of the DL response with stimuli outside the visual field of the subject, its results also allow a preliminary analysis of localization accuracy for a source near the head. Figure 7 shows the relationship between the azimuth of the sound source and the azimuth of the subject response on each trial. The solid line shows the "correct" responses. Extreme front-back reversals are seen in the responses at the top left and bottom right of the plot. Stimuli at all elevations are included, so some poor azimuth responses may be a result of a high elevation. The most notable feature of the azimuth responses is that the spread of responses is much smaller around 0° than around $-90°$. This is consistent with previous localization studies which have shown that localization accuracy in azimuth is better for sources in front of the listener than sources at the listener's side.

Figure 8 is similar to Figure 7, except the the stimulus and response elevations are plotted instead of the stimulus and response azimuths. There is clearly a strong correlation between the stimulus and response, but there are no unusual or interesting features in this plot.

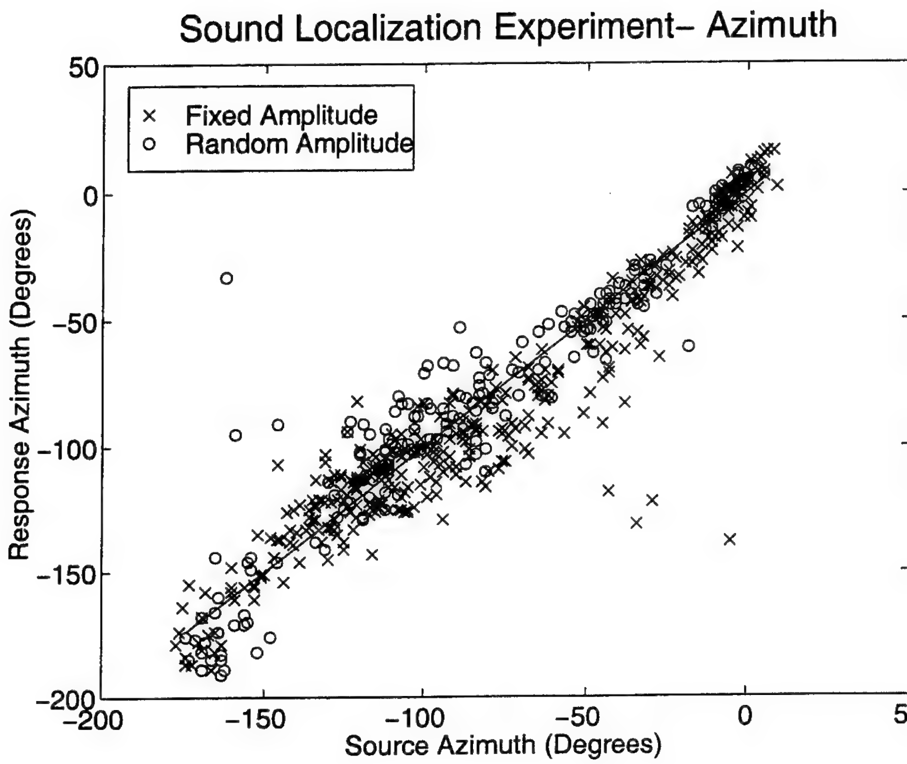

Figure 7: Stimulus azimuth vs. response azimuth in psychoacoustic experiment

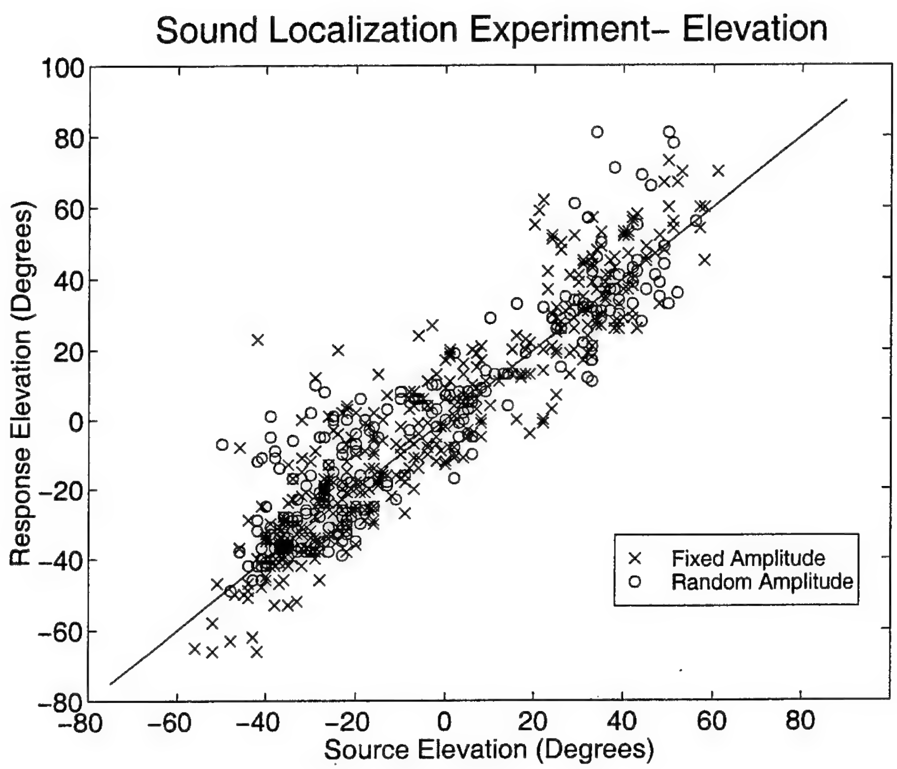

Figure 8: Stimulus elevation vs. response elevation in psychoacoustic experiment

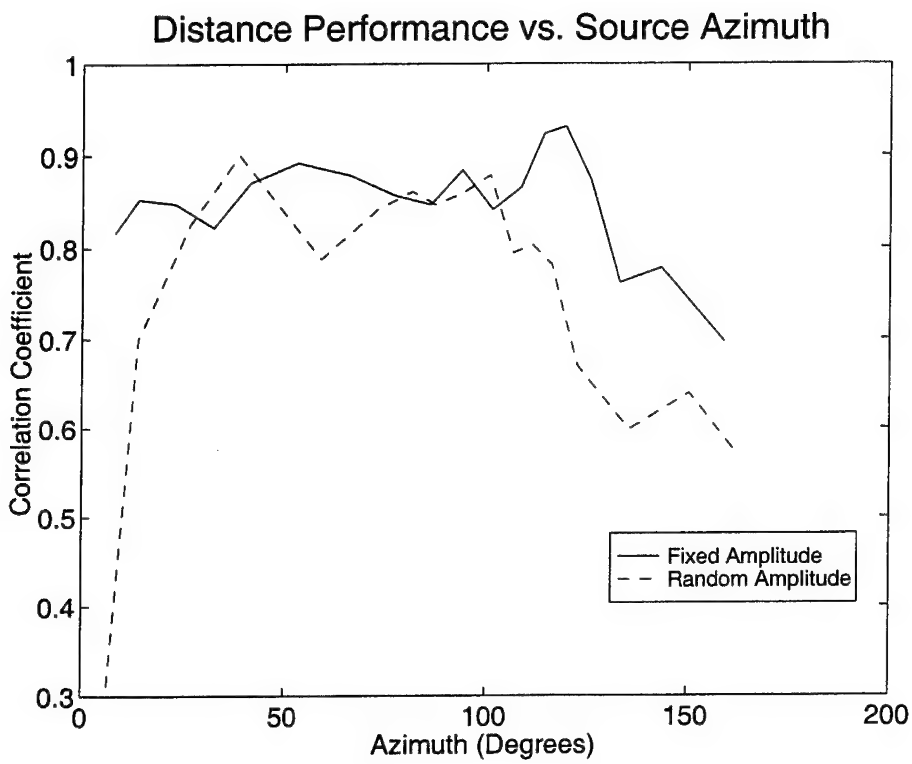

Figure 9: Correlation between log stimulus distance and log response distance vs. azimuth

By far the most striking result of the psychoacoustic experiment is shown in Figure 9. This figure shows the relationship between audio distance perception and source azimuth. The trials were first sorted by the absolute value of the azimuth, so the small number of trials on the left side of the listener were reflected across the median plane and combined with trials on the right side. Then the sorted trials were separated into 19 overlapping bins, each containing approximately 40 trials. In each bin, the correlation coefficient between the log of the source distance and the log of the response distance was calculated. Finally, the correlation coefficient in each bin was plotted as a function of the mean azimuth value in each bin. The resulting graph summarizes the relationship between stimulus and response distance as a function of azimuth. The correlation coefficient for the fixed amplitude experiment is approximately 0.85 at all azimuth values from 0° to 135°. In the random amplitude experiment, the correlation coefficient is also approximately 0.85 for azimuths greater than 30°, but falls off rapidly to 0.32 as sources move closer to 0°. Both experiments show a drop-off in performance at azimuths greater than 135°, but this may be a result of reduced accuracy in the pointing response for locations behind the subject.

These data indicate that audio depth perception is roughly independent of azimuth when the amplitude of the source is fixed, but that it degrades substantially in the median plane when the source amplitude is randomized. The results suggest that binaural distance cues dominate in the near-field when amplitude cues are not available, as depth perception was worst in the median plane where binaural cues disappear. Note, however, that the data from this experiment are preliminary, and more experments are required before firm statements can be made about auditory localization for sources near the head.

5.2.2 Comparison of performance in front and rear hemispheres

Results were analyzed separately for stimuli in the front and rear hemispheres of the subject (trials with azimuth $> -90°$ and trials with azimuth $< -90°$). The trials were about evenly divided between the front and rear (208 front, 184 rear). The results can be summarized as follows (see Table 13):

- In the front hemisphere, the errors with the DL response and an acoustic stimulus are much larger than the corresponding DL errors seen earlier with a visual stimulus. Specifically, the angle error of 12° and vector length error of 37% in Experiment 3 are both approximately double the values in Experiment 1 (6° and 21%, Table 2). For the

SH method, however, the errors are approximately the same in Experiments 1 and 3. These results suggest that the response error due to uncertainty about the location of the auditory target dominates the DL errors, but the inaccuracies inherent to the SH response method contribute substantially to its overall variability.

- The vector-length and angle errors with the DL method and an auditory target (Experiment 3) are both comparable to the best performance seen in the SH response with a visual target (Experiment 2, see Table 10). In other words, the combination of variability associated with auditory localization and the noise in the DL task is about the same as the noise in the SH response alone.

- For each response method, performance in the front and rear hemispheres is similar, both in terms of mean errors and standard deviations. Although a few error measures do increase significantly (SH angle error is 22% higher in the rear hemisphere), overall differences are small. This result is surprising. *A priori* one would expect the DL task to be substantially less accurate when the subject is forced to respond outside the visual field, and at locations in the rear hemisphere that are relatively difficult to reach. However, these data indicate that average performance is virtually identical in the front and rear hemispheres. Thus, any degradation associated with positioning the response sensor out of the visual field is small compared to the overall noisiness of the auditory localization task.

6.0 CONCLUSIONS

- The Direct-Location (DL) response is markedly superior to any of the alternative response methods tested. When the DL response was the first response made with a visual target, the mean angle error with DL was only 4° and the mean vector-length error was only 12%. In contrast, the next-best response method produced angle and vector-length errors roughly three times as large. DL was also the only response without significant biases in azimuth, elevation, and distance. Clearly the DL response is the most appropriate for a near-field localization experiment.

- Replacing the visual target with an acoustic source approximately doubled the errors in the DL response method. The errors in the acoustic experiment were, however, approximately equal in the front and rear hemispheres. Although it is possible that the DL response was noisier in the rear hemisphere than in the front hemisphere, it appears that the DL response errors in either hemisphere are small compared to the uncertainty in the auditory localization task.

- Indirect location with a large or small manikin head (LH and SH) and the verbal report of coordinates (VR) were all similar in performance once the responses were corrected for overall biases. Angle errors ranged from 13.2° to 16.5°, and vector-length errors ranged from 35% to 39%. Although VR was slightly more accurate, it was slower than the other response methods. The SH response was always at least as accurate as the LH response, indicating that scaling distance down by a factor of two did not adversely influence response accuracy.

- Although variations in response biases across subjects were evident, bias trends common to all subjects were substantial. In particular, a bias correction that depended only on response elevation improved performance at least 60% as much as a more complicated correction scheme based on the individual subject biases.

- When giving all four responses on each trial (Experiment 1), response accuracy was affected by response order only for DL, and not for any of the three "indirect" response methods (LH, SH, and VR). The DL errors were about twice as large when it was the last response than when it was the first response.

- Response errors in the LH and SH methods were highly correlated, implying that both responses may be derived from the same encoded memory of the target location. Correlations between the other response methods were small.

- Once the responses were bias corrected, the errors were roughly independent of target location. Vector-length error was negatively correlated with distance for all four response methods, but the correlation was relatively weak (average $r=-0.28$).

- The SH response was not significantly different in angular error when the manikin head was profiled than when it faced the same direction as the subject.

In summary, the DL was superior to the three indirect response methods in terms of having the best overall accuracy, the smallest biases, and the shortest response time. These advantages promote DL as an attractive response metric for future localization studies. One caveat to this recommendation is that the present work utilized stimulus locations restricted to one side of the head. This restriction allowed positioning of the response sensor with a wand without the need for passing the wand from one hand to another. Such passing or some other modification might be required if stimulus distances up to a meter at any angle about a subject were to be included for testing. Evaluations of such response modifications await future research.

7.0 REFERENCES

Brungart, D. (1998). *Near-Field Auditory Localization.* Ph.D. thesis, Massachusetts Institute of Technology.

Butler, R., Levy, E., & Neff, W. (1980). Apparent distance of sounds recorded in echoic and anechoic chambers. *Journal of Experimental Psychology, 6*, 745–750.

Coleman, P. (1968). Dual role of frequency spectrum in determination of auditory distance. *Journal of the Acoustical Society of America, 44*, 631–632.

Durlach, N., & Braida, L. (1969). Intensity perception. i. preliminary theory of intensity perception. *Journal of the Acoustical Society of America, 46*(2), 372–383.

Gilkey, R., Good, M., Ericson, M., Brinkman, J., & Stewart, J. M. (1995). A pointing technique for rapidly collecting localization responses in auditory research. *Behavioral Research Methods, Instrumentation & Computers, 27*, 1–11.

Makous, J., & Middlebrooks, J. (1990). Two-dimensional sound localization by human listeners. *Journal of the Acoustical Society of America, 87*, 2188–2200.

Mershon, D., & Bowers, J. (1979). Absolute and relative cues for the auditory perception of egocentric distance. *Perception, 8*, 311–322.

Soechting, J., & Flanders, M. (1989). Sensorimotor representations for pointing to targets in three-dimensional space. *Journal of Neurophysiology, 63*, 582–592.

Wightman, F., & Kistler, D. (1989). Headphone simulation of free-field listening. ii: Psychological validation. *Journal of the Acoustical Society of America, 85*, 868–878.

Wightman, F., & Kistler, D. (1992). The dominant role of low-frequency interaural time differences in sound localization. *Journal of the Acoustical Society of America, 91*, 1648–1660.

APPENDIX A: COORDINATE SYSTEM DEFINITIONS

One of the key goals of this research is a comparison between judgments of location relative to the subject's own head and judgments of location relative to a manikin head. This requires a coordinate system which is consistent for both the subject's frame of reference and the manikin head's frame of reference. Three features common to both the human and manikin head were used as the basis of this coordinate system: the openings of the left and right ear canals, and the tip of the nose. The locations of these three features generate a coordinate system for each head centered at the midpoint of the interaural axis and with a horizontal plane approximately parallel with the floor. Specifically, head-referenced Cartesian axes were defined as follows: The Y axis of the coordinate system is the interaural axis of the head, and is positive on the left side. The X axis is the perpendicular bisector of the interaural axis passing closest to the tip of the nose, and is positive for locations in front of the head. The Z axis is perpendicular to both the X axis and Y axis and is positive above the head.

The reference points for these coordinate systems were measured with the 3-Space tracker before each block of trials. The tracker provided the X, Y, and Z locations of each position relative to the 3-Space source, which was mounted on the chin-rest. The locations of the left and right ears and the tip of the nose, represented as column vectors $\vec{E_l}, \vec{E_r}$, and $\vec{N}$, were measured for the subject, the large manikin head, and the small manikin head with a 3-Space Tracker sensor. These locations were used to determine the origin and the directional cosines of the X, Y, and Z axes for each of the three heads.

The origin of the coordinate system $\vec{O}$ is defined as the midpoint of the interaural axis, $\frac{\vec{E_l}+\vec{E_r}}{2}$. The interaural axis also defines the directional cosines of the Y axis, $\vec{D_Y} = \frac{\vec{E_l}-\vec{E_r}}{|\vec{E_l}-\vec{E_r}|}$. The X axis is defined by the tip of the nose, the origin $\vec{O}$, and the directional cosines of the Y axis $\vec{D_Y}$. Ideally, the X axis should pass through the tip of the nose, but the vector $\vec{N} - \vec{O}$ is not necessarily perpendicular to the Y axis, as seen in Figure 10. When this is true, the projection of $\vec{N} - \vec{O}$ onto the Y axis is subtracted from $\vec{N} - \vec{O}$ to determine the perpendicular bisector of the interaural axis which is closest to the tip of the nose. If we call

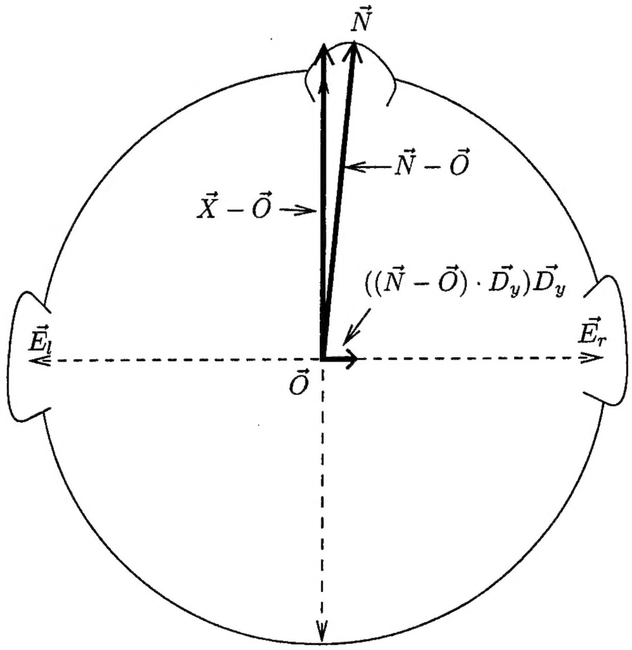

Figure 10. Definition of the X-axis

An example where the vector $\vec{N}-\vec{O}$ is not perpendicular to the interaural axis. The projection of $\vec{N}-\vec{O}$ onto the Y axis, $(((\vec{N}-\vec{O})\cdot\vec{D_Y})\vec{D_Y})$, is subtracted from $\vec{N}-\vec{O}$ to yield the vector $\vec{X}-\vec{O}$, which is a perpendicular bisector of the interaural axis. Normalization of $\vec{X}-\vec{O}$ gives the directional cosines of the X axis.

this vector $\vec{X} - \vec{O}$,

$$\vec{X} - \vec{O} = (\vec{N} - \vec{O}) - ((\vec{N} - \vec{O}) \cdot \vec{D_Y})\vec{D_Y},$$

and the directional cosines of the X axis are found by normalization:

$$\vec{D_X} = \frac{\vec{X} - \vec{O}}{|\vec{X} - \vec{O}|}.$$

Finally, the directional cosines of the Z axis are obtained by taking the cross product of the $\vec{D_X}$ and $\vec{D_Y}$ vectors, so $\vec{D_Z} = \vec{D_X} \times \vec{D_Y}$.

The directional cosines can be used to convert the Cartesian coordinates of any location, measured relative to the head tracker, into the transformed coordinate system relative to the head. Let $\vec{S}$ be an XYZ location relative to the source of the 3-Space tracker. First the XYZ coordinates are moved relative to the center of the head by subtracting the origin of the transformed coordinate system $\vec{O}$ from $\vec{S}$. Then the XYZ coordinates of the vector are projected onto the X, Y, and Z axes of the transformed coordinate system by matrix multiplication

$$\vec{S_T} = (\vec{S} - \vec{O})' \cdot [\ \vec{D_X} \quad \vec{D_Y} \quad \vec{D_Z}\ ].$$

where $'$ denotes matrix transposition. The resulting vector $\vec{S_T}$ can be used to determine the azimuth, elevation, and distance of the target relative to the coordinate system defined by one of the three heads, thereby allowing a direct comparison of all four response methods.